DOCUMENT RESUME

ED 201 706 UD 021 416

AUTHOR Jones, Leon
TITLE Educational Evaluation: Service or Menace.
PUB DATE [79]
NOTE 26p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Educational Assessment; Elementary Secondary Education; *Evaluation Methods; Evaluative Thinking; *Formative Evaluation; *Program Evaluation; *Summative Evaluation
IDENTIFIERS Provus Discrepancy Evaluation Model

ABSTRACT

This paper analyzes educational evaluation and assesses its value to the educational process. The development of evaluation research and practice is outlined and the paper focuses on a comparison of formative and summative evaluation techniques. Both of these types of evaluation are discussed with an emphasis on formative evaluation. The facets of formative evaluation examined include: (1) its nature and application; (2) how it is conducted; (3) when it should be implemented; (4) where it is conducted; and (5) who implements the process. In order to explore the extent of educational service provided by evaluation, the Malcolm Provus Discrepancy Evaluation Model (DEM) is reviewed. A general analysis of evaluation with special attention to the possibilities for negative effects concludes the paper. (APM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED201706

Educational Evaluation:

Service or Menace

by Leon Jones

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Leon Jones

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SP021416

MAR 2 6 198

EDUCATIONAL EVALUATION: SERVICE OR MENACE

Educational evaluation got its real break-through during the mid- to late-1960's. The passage of Title I of the Civil Rights Act of the 1964 was a real boon to the evaluation of educational programs. From a theoretical standpoint, in terms of an academic discipline, the evaluation aspect of Title I gave rise to the present day mission and function of program evaluation. This mission and function have been reduced to the following definition of program evaluation which prevails throughout this paper. In 1968, Thomas Hutchinson[1] defined evaluation in terms of its fundamental purpose; namely, a systematic procedure for gathering and analyzing data so as to provide useful information for decision-making by decision-makers. The ethos of this definition has been stated variously by other prominent evaluators and evaluation researchers dating back to the period of the late-1960's and continuing up through the present time. They include Daninel Stufflebeam,[2] Malcolm Provus,[3] and more contemporaneously Peter Rossi, Howard Freeman, and Sonia Wright.[4]

Prior ro the end of the decade of the 1960's, in answer to the question: what's in a name? one prominent evaluator posited that there is much to be said for what's in a name or label. C. Robert Pace[5] points out that there is a wide assortment of labels that intimate evaluation activities. His thesis is that many of the things bearing the evaluation label could and, indeed, should be called by other names or labels. His position is that the way we label an activity influences the way we see that activity's role and function. It is this anticipated conceptual framework that makes the writers worry about the uncontrolled smorgasbord[6] approach to the labeling of activities as "evaluation."

Assuming arguendo that the careful labeling of an activity as "evaluation" is within reach and that from a practical point of view the "evaluation" label refers to ongoing program activity that is amenable to a systematic evaluative approach, the issues of service and menance have relevance. If, on the one hand, evaluation of a program means that pertinent data about that program will be converted into useful decisions about that program, then it appears reasonable to characterize the evaluation experience as a service. This is especially so for those decision-makers that eventually will use that data. If, on the other hand, as Provus so elegantly articulated wherein he states that:

> To the school administrator who needs information about the effectiveness of school programs, the word "evaluation" conjures up some unpleasant memories: a report that took "too long" to prepare and overlooked the obvious while concentrating on the trite, a university consultant who proved unintelligible and eventually hostile, an investigator who got in everyone's way and never seemed able to draw definitive conclusions.[7]

We discuss these service and menance aspects in more detail in a section that follows.

In the meantime a few words about the state of the art of evaluation appear to be in order. As intimated earlier, the boon that Title I of the ESEA programs gave to evaluation as an educational discipline carried with it a number of problems, the most notable one being that evaluation work preceded evaluation training. The matter of having evaluation training follow the manifestation of numerous evaluation jobs is just the opposite of the typical relationship between academic training and the work force. The full impact of ills of the magnitude of this reverse approach to program evaluation as an academic discipline has probably not yet run its course. In any event, by the mid-1970's Daniel Stufflebeam made this assessment:

"If traditional evaluation efforts are anything they are noncredible."[8] In effect, this means that present day evaluation training possibly tracks the training efforts of more than a decade ago. At that time evaluation training was not only highly personalized, but it rested largely with the founders of a few evaluation models such as Stufflebeam's own Context, Input, Process, and Product (CIPP) model,[9] Stakes' "Countenance of Educational Evaluation model,"[10] and Provus' "Discrepancy Evaluation Model,"[11] to name a few.

Because evaluation work at the field level antedated evaluation as a theoretical or conceptual discipline, it appears certain that many evaluation practitioners and consumers alike have never had as much as one credit hour of academic training in evaluation. It is well established, for instance, that colleges and universities have never offered courses or training adequate to meet the evaluation demand.[12] Even though program evaluation is well into its second decade of experience, not only does the evaluation theory continue to lag behind the evaluation demand, but seemingly too few of the nations academic institutions have come forward to train evaluators in sufficient numbers.[13] In the meantime, both the evaluation activities (by untrained evaluators) and the evaluation demands seem to continue to enjoy a steady and momentous increase. Ironically, where evaluation research and methodological activity appear to be under the way,[14] their foci tend to be narrow and of somewhat parochial aims at audiences that are small and isolated that tend to be of limited practical value in terms of the diversity needed at the field level.

Prior to the advent of the 1980's, fruits of the evaluator's labor were beginning to enter the market-place in the form of books, monograph series, and convention reports via ERIC Documents[15] and

cassette recordings. As recent as 1979, however, brochures and announcements about the book: Evaluation A Systematic Approach[16] proclaimed that "The wait was over!" This book was hailed as being the first textbook on program evaluation. Unlike the parochial oriented models and schemata proffered by Stufflebeam and other psychologists, behavioral objective specialists, and philosopher, this books' authors and social scientists Rossi, Freeman, and Wright appear more global in their orientation. They apppear to be steeped in the training, evaluation, and program management of developing countries at the international level. Therein might possibly lie the books' greatest shortcoming, however. Whereas the typical program evaluation model tends to be either conceptually inadequate or so patently complex until neither its creator nor his most ardent student can apply it fully at the field level on a given program; this book envisions, among other things, the wherewithal to evaluate multi-faceted programs on the order of a typical developing country. What appears to be needed more than anything else regarding the evaluation of educational programs is a viable methodology capable of evaluating a typical educational program that has a set of ascertainable goals or objectives. It would then seem to be no very strenuous normative leap for either the evaluation practitioner or consumer to draw the conclusion that that methodology could be applied to the lot of educational programs generally.

Most of the evaluation of educational programs to date come under the rubric of two types of evaluation: formative and summative. We examine those types of evaluation in the section that follows.

Summative Evaluation vs. Formative Evaluation

As mentioned earlier, shortly after the Federal Government's

generous financial support for evaluation via its ESEA programs of 1965, a burgeoining evaluation initiative came into being. That initiative, though plagued with sustained growing pains, still obtains. An outcome of early evaluation initiative has been the product of Michal Scriven's extension of the foundation laid by Lee J. Cronbach[17] regarding distinctions and procedures that have proved to be of great practical value to the evaluation practitioner. The valuable end results of Scriven's work was the emergence of two formidable types of evaluation, namely, formative and summative evaluation. Formative evaluation and summative evaluation are mutually exclusive and they continue to enjoy present day wide usage. Before going into their usage proper, we first define these evaluation concepts. In defining formative and summative evaluation we draw upon William L. Goodman's expressed position which is that:

> Formative evaluation is directed toward ongoing activities or products currently being made, evaluating their merit, and feeding data back into the cycle to reform the activities or product. In effect, judgments are repeatedly being made about a product or activity as it is being formed or developed.
>
> Summative evaluation relates more to a judgment of the overall worth of an activity or product, normally at an end or critical decision point. It often takes the form of a comparative evaluation.[18]

A cursory glance at Goodwin's definition of formative and summative evaluation tends to indicate that the clue to Scriven's use of the terms "formative" and "summative" lies in their respective root meanings. A somewhat overly simplistic representation of the relationship between summative and formative evaluation may be characterized as follows.

In brief, summative evaluation generates end-of-program data aimed at executive management personnel, funding sources, and other outside audience persons. The evaluation data translates into infor-

mation that speaks to whether or not the program was successful with respect to goal attainment. Whereas an affirmative answer to this question often leads to refunding or continuation initiative; a negative reply to the question of goal attainment can place the continuation of a program in jeopardy.

In the real world of evaluation the matter of accomplishing goals is generally not so weighty as a single issue. Quite often, the single most weighty issue is: of what practical use is the program outcome irrespective of the evaluation results? Here the actual benefit of the program can be balanced against the harm of its failure to attain specific goals vis-a-vis the evaluation. Although program goals are extremely important in terms of quality assurance and cost effectiveness, the indicated balance approach may prove to be a critical factor in a decision on the termination of a program. The nature of on-going program activity is such that a good faith effort to achieve pertinent goals can lead to a hefty partial goal fulfillment of so much magnitude and a desirable outcome, to boot, until reasonable persons might tend to agree that support for the continuation of the program is the only logical decision to be made --the failure to complete portions of some goals notwithstanding.

Formative evaluation plays an integral part in the outcome of program activity even though the goals of the program may be only partially attained. It goes without saying, perhaps, that a program's successful attainment of its goals usually involves a dependence on the work of the formative evaluator. This involvement can be explained in terms of the relationship between formative and summative evaluation.

We begin by looking at the purposes of each type of evaluation. Formative evaluation, for example, is "primarily directed

toward on-going activity or products currently being made, evaluating their merit and feeding data back into cycle to reform the activities or products."[19] The primary purpose of formative evaluation is thus the developmental improvement of an activity or a product. In the alternative, summative evaluation is implemented after development of the activity has been completed or at a critical decision point, as explained above. Thus, the primary objective of summative evaluation is the determination of the overall worth of the activity or product.

Within the context of these two completely different purposes, one factor that seems to stand out as a distinguishing feature is "change." "Change" of program elements is an integral feature of the formative evaluation process. It is this factor that makes the formative evaluators role distinguishable from that of the summative evaluator who goes to great lengths to hold the program or product constant.[20] Guarding against changes in summative evaluation appears to be necessitated by the fact that the evaluator seeks to attribute the outcome or product of the program as its end-of-program results. Therefore, any changes made in the program while conducting summative evaluation will make it difficult to draw conclusions about the effectiveness of the program. Summative evaluation has to unambiguously state the effect of the program or product in competition with another. The results of the summative evaluation are of primary concern to those financiers who will eventually decided whether or not the program undergoing evaluation is to be continued or not. As a side benefit, summative evaluation provides the basis for policy decisions that do not necessarily concern revision of the program or product.

The need to hold program goals constant means that there has

to be a close monitoring of the program to insure that the program is implemented as per the dictates of pertinent goals. This monitoring is actually a formative aspect of evaluation which means that herein lies a feature of formative evaluation that is closely linked to summative evaluation regarding its constancy criterion.

Be it summative evaluation or formative evaluation, quite often, the successful outcome of a program activity and the nature of the work of the evaluator appear to be somewhat less than an arms-length apart. This means the work of the evaluator possibly tends to influence to outcome of the program activity. More importantly, it means that the evaluator must minimize that subjectivity with an adequate dosage of objectivity in his evaluative approach in its stead. There may be times when the nature of evaluation carries with it an element of threat to the program manager incident to his day to day operations of the program activity undergoing the evaluation. This factor combined with the possible subjective influence that the evaluator could have on the influence of the outcome of an evaluation depict some inherent ills of the evaluation modus operandi per se. These inherent ills seem to be a particularly challenging aspect of the work of the formative evaluator. It is in the spirit of formative evaluation that gives rise to the nature of the section that follows.

The Nature and Application of Formative Evaluation

Now that the involvement of evaluation in educational activities has expanded beyond curricula activities to include social action programs, there is a pressing need for the application of bold new and revised evaluation concepts. It is the evaluation of social action programs that seems to have provided the practitioner

his greatest challenge. Suffice it to say that the scope of program activity is widespread. And, yet, the skill and training of the evaluator continue to be called into question because the demand for skilled evaluators has remained out of line with the supply of well trained (and oftentimes self-proclaimed) evaluators.[21]

Up through the present time, practitioners have suffered from the absence of formal training in evaluation. They have received their training in evaluation techniques on a piecemeal basis for the most part.[22] Some of the early practitioners, who were able to find evaluation training, received courses in test and measurement; some of their training were possibly acquired in educational statistics; and other parts of their training were tacked on to methods courses perhaps.[23]

We intimated earlier that the evaluation jobs and, hence, the need for the evaluator preceded both the availability of evaluation trainees and evaluation courses for the training of evaluators to meet the evident evaluation needs or upcoming jobs in evaluation.[24] In a way, it seems somewhat unexplanable and rather unfortunate that after nearly fifteen years of history, to a measurable extent, the initial gap between the supply and demand of trained evaluators remains unabridged. As early as 1972, this writer addressed one aspect of the gap facing the evaluation practitioner.[25] Some three years later, Jerome Sanders[26] complained that there were only three institutions of higher education with meaningful evaluation programs. As of this writing, there appears to be no appreciable evidence of a change in the state of affairs as expressed by Sanders in 1975.

Presumably, the conspicuous absence of trained evaluators serving at the field level and serving as practitioners generally has played a profound part in crystalizing the negative image that

tends to permeate the range and domain of program activity. Admittedly there are ambiguities about the theoretical concepts of evaluation as well as the misapplication of various theories at the field level. These shortcomings not withstanding, increasingly program directors and other high-level persons responsible for program development and implementation have begun to see the value in systematic evaluation, as can be attested by the facts that books on the topic are now appearing in the market-place. See note 16 above. Having identified an audience for program evaluation work, we now turn our attention the nature of evaluation proper by exploring the questions that follow.

How is Formative Evaluation Conducted or Carried out? In addition to the foregoing concepts proffered about formative evaluation, evaluation concerned with in-progress program improvement is called formative evaluation. It is directed towards on-going activities of the program under consideration. In effect, judgments are to be made cyclically or on an interim basis about the program activity as it is being conducted or developed.

For purposes of this paper, formative evaluation is considered to be goal oriented evaluation (as opposed to goal-free evaluation). So then how we go about conducting evaluation depends on an evaluation that is capable of accommodating an arbitrary program. This factor aside for the moment, we explore an evaluation approach that contains both generalizability across programs and goal analysis within programs in terms of the evaluation scheme. The evaluation scheme that influences this discussion consists of three broad categories. They are the "the beginning stage of the evaluation process," "the evaluation modus operandi," and "the summary of the evaluation."

During the beginning stage it is useful for the evaluator to initiate a planning conference. This conference should involve the evaluator, decision-maker, decision-maker designee or program manager, etc. Here the basic evaluation model is reviewed; the commitment to the evaluation regarding the decision-maker's cooperation, participation, and authority are explored; the day-to-day operations of the evaluative process such as evaluator's access to the decision-maker, handling of interim or cyclic evaluation reports etc. are discussed. It is during this stage of the evaluation that the evaluator normally commits himself to preserve the decision-maker's goal intentions and by doing so facilitate the ever presence of a distinction between the functions of the decision-maker and the evaluator.

Among the important aspects of the evaluation modus operandi under consideration are the following aspects: (a) statements of goals which consists of the acquisition and identification of pertinent goals, (b) evaluation design which calls for the operationalization of goals,[28] (c) evaluation procedures which calls for the development or acquisition of instruments or assessment techniques and the implementation of those materials, (d) data collection and analysis, and (e) a report of the evaluation outcome.

These few words are merely and overview of the evaluative process and they should be interpreted in that light. Otherwise, they might create an oversimplication of the evaluation process in the real world. An overview about another aspect of formative appears in the subsection that follows.

When should Formative Evaluation be Implemented? We explore this question by first explaining the point of interest that it

raises. Our interest is knowing when the optimum time for the formative evaluation to be conducted really is from an evaluation perspective. Careful search of evaluation literature highlights the fact that the theoreticians and researchers have not yet articulated a common point in time for the formative evaluator to begin his work.

But, nevertheless, there appears to be wide support for the evaluator to begin his work at the beginning of program activity under consideration. The bulk of this support can be found in the basic nature of the early evaluation models that have been created by various evaluation researchers. Trends of this effect can be found in "The Process of Decision-Making And Its Relationship to Contemporary Education Models."[29] The work of the evaluation researchers contain therein make it appear reasonable to conclude that the optimum time for the evaluation starting point is at the beginning of the program's activity.

Because the goals undergird so much of the evaluation process in terms of what must follow their acquisition, identification, and operationalization, it seems fitting that evaluator and the decision-maker begin their work simultaneously. This suggested time frame seems especially fitting as we reflect on the mattter of securing and procuring goals and the importance their function relative to the evaluative process. The matter of securing and procuring goals receive some attention in the subsection that follow.

Where is Formative Evaluation Carried Out? Because of the nature of aspects of the "evaluation modus operandi" as mentioned in an earlier subsection, formative evaluation may be carried out in several different ways and at different locations simultaneously. For example, consider the wide reach of goal identification. Goal statements may occur in pertinent written documents or oral remarks.

In the main, however, formative evaluation is carried out at the field level.

As intimated earlier, usually the formative evaluator works closely with the program manager regarding his day-to-day operations and, hence, he normally works at the program site . The nature of the work of the formative evaluator, in terms of operationalizing goal statements, normally involves transforming somewhat vague statements into either behavioaral statements or behavioral states and can usually be best carried out at the program site.

Because the operationalization of goals at the field site usually includes the involvement of program staff persons, the operationalization of goals tends to pave the way for implementing the procedure which follows. These procedures include development or acquisition of instruments or assessment techniques. The data collection aspect then follows. It too, must be conducted at the field level. Beyond this point in the evaluative process, other features such as the analysis of data and the summary or report of findings need not be conducted at the evaluation site. Although these features can be conducted away from the program site, they are sometimes conducted on site as a means of convenience.

In view of the foregoing, it can be said that, for the most part, formative evaluation is carried out at the field level. For additional information about the implementation of formative evaluation, we explore the question that follows.

Who Implements Formative Evaluation? Before answering this question proper, the following analogy has some relevance. So we leap to the issue of fighting a world war. The question then becomes: Who fights the war? Is it the Commander-in-Chief, or is it

the field soldier? Without pondering over these questions we shift to evaluation, and, as we think about the question: Who implements formative evaluation, then it seems reasonable for us to conclude that there is no simplistic, direct answer. First we ask is it the evaluator himself? Next, we might ask is it the evaluator's field staff? Since there appears to be no unique answer to the first question, is it then reasonable to conclude that the answer lies in a combination of both questions? The writers are of the opinion that it all depends; that is, it depends on some other factors.

The principal factor regarding who implements the evaluation depends, in great measure on whether the entire evaluation involves more persons than the solo evaluator. In such cases, then clearly it is the evaluator who implements the formative evaluation. When the evaluation activity consists of more than the solo evaluator, the question: Who implements the evaluation usually refers to the principal evaluator. This is so whether the principal evaluator does any part of the actual evaluation tasks or not. Typically, the evaluator uses non-evaluators to collect the evaluation data, but he then conducts the analysis of the data himself. William Goodman intimates what the role of the person who does the evaluation might be like in the characteristic of formative evaluation[30] that follows.

CHARACTERISTIC	EVALUATION TASK
Primary purpose	Developmental improvement of activity or product
When used	Continually as it provides feedback into developmental cycle
Investigative style	Rigorous, systematic, diagnostic
Personnel involved	Internal staff or internal program evaluators
Users of information collected	Program staff, developers, and insiders.

SERVICE OR MENACE

"To err is human" as part of this familiar adage goes. To glory in the shortcomings of others appears to be only too common in both our academic community and society in general. As an emerging academic discipline, educational evaluation has been unable to avoid its share of criticism and attacks on its credibility. In fact, the call for improvement of its credibility by members from its own ranks have now become somewhat dated.[31] As an academic discipline, it is questionable as to whether evaluation; which is highly stylized because of the influence that prominent persons have given to its development, can pass academic scrutiny. Moreover, we wonder whether the influence of the notable personalities on evaluation's development has made it too parochial or unresponsive to consumer need. And further, because of profit oriented aspects of evaluation, it is questionable as to whether this feature has been dispositive regarding its acceptance within the academic community. These factors in concert with the charge of the basic definition of evaluation: "[t]o produce decision-making data by decision makers,"[32] tend to lay the ground work for an exploration of the question whether evaluation of educational programs provides a service or menace to its consumers. First we look at the service aspect.

Service

To explore the "service" aspect proper we pick one of the existing evaluation models that has some history with program evaluation at the field level; namely, Malcolm Provus' Discrepancy Evaluation Model (DEM). Basically, the DEM establishes the standard in developing the program design. In practice, this standard is used for purposes of comparing the program performance. A deviation from the standard in prog-

ram performance connotes a discrepancy. If a discrepancy does exist, a decision must be made with reference to program improvements. To facilitate understanding how the DEM works, we explain the use of the following terms contextually.

Standard: a criterion established by program developers used to compare performance as appropriate;

Performance: measurable program activity in behavioral form that is called for by the program design; and

Discrepancy: inconsistency or difference between program performance and the standard.

Having highlighted some critical terms of the DEM,[33] we now focus on the service aspect of evaluation per se. Elements of the DEM will serve as a basis to show that evaluation purports to provide service to intended consumers via the program design, whose sole function is to generate measurable program outcome. It is noteworthy that in the educational arena the concept of "consumer" normally spans the gamut in terms of interest of the State via its Department of Education, the Local Educational Agency, the public school and the family. It goes without saying, perhaps, that the service interest of these different factions not only differ markedly but they may even conflict. Thus, based upon the goals or the dictates of the program intents, the task facing the evaluator is to develop a viable program design that accommodates these varying interests on the one hand and, at the same time, produce outcomes tailored to the varying interests on the other hand.

The effectiveness of the DEM, with respect to the service issue, is a function of the completeness of the standard against which the performance can be measured. Herein possibly lies the DEM's strongest and weakest points as the evaluator and the decision-maker strive to carry out independent functions at the program site concurrently. This is so because on the one hand the evaluator must strive to refrain from

disrupting the program manager either physically, sociologically, or psychologically. And yet, on the other hand, inherent in the nature of the work of the evaluator as he establishes the program design is his responsibility to consult with the program manager personally and on a sustained basis; i.e., throughout the development of the program design. It is beyond the scope of this paper to treat "program design"[34] proper. Briefly, however, we mention that the essence of the program design is the effective use of its two aspects: its IPO's (inputs, process, & outputs) and its various levels (I, II, & III) of design network.

To fully appreciate the service aspect of the DEM we look at its entire evaluative process, which includes the following stages: (a) design evaluation, (b) input evaluation, (c) process evaluation, (d) output evaluation, and (e) cost-benefit analysis. It is beyond the scope of this paper for us to treat these stages of the DEM, rather, we note that the merits of service per se more appropriately attaches to successful completion of "the evaluation cycle" which consists of the following five parts in the order given: (a) develop program design, (b) plan evaluation, (c) collect performance information, (d) report discrepancy, and (e) resolve discrepancy.[35] From within this evaluation cycle the matter of resolving the discrepancy is critical as the notion of evaluation, as a service, attaches great weight to the ultimate outcome. Arguably, if the discrepancy is resolved then the ultimate outcome will be the manifestation of the intended goal. Actual realization of this desired state of service depends on the activity at the field level and must be observed on a program by program basis. We now turn our attention to the menace aspect.

Menace

As can be gleaned from an earlier section of this paper evaluation as a concept connotes the assessment of the worth or value of a program.[36] In a like manner it is therefore reasonable to evaluate the evaluation of a program under consideration. Assuming arguendo, that an effort is made to assess the worth or value of the evaluation, then one approach could be as follows. Consider the evaluation operationally at the field level to determine whether there exist strands or traces of its manifest purpose; query, was there evidence of a systematic procedure geared specifically to producing decision-making data for decision-makers. Since we are interested in formative evaluation then first we look for the systematic procedure and next we examine the functional aspects of that systematic procedure at the field level. As we examine the functional aspect of the formative evaluative process an in-depth assessment of the inner-working features that characterize the preceding section of this paper would be required.

The principal concern of the assessment of the evaluative process reduces to an examination of the role of the evaluation throughout the various aspects of the evaluative process with an eye toward identifying those elements of the evaluation that can be labeled as "menace." Here menace enjoys the wide range of operational concepts. In the world of evaluation "menace" is normally thought of as comprising the negative aspects of the evaluator / decision-maker joint ventures in general and other non-specific evaluator related experiences from time to time occuring in particular that are negative and, hence, constitutes a menace to all parties involved.

Essentially three issues of typical concerns surface upon considering the menace aspect of evaluation. They are: (a) how the program activities comport with the description and definition of goal intentions; how the observations conducted during the data collection phase constitutes a menace; and how the interpersonal relationship between the evaluation staff and the program staff constitute a menace. Inherent in the purpose and definition of evaluation is the use of decision-making data by decision-makers.[37] Normally data generated as a result of the evaluative process and used to make decisions by decision-makers are more likely to be accepted when the data generated is a function evaluator/decision-maker team work. Presumably the team work between the evaluation staff and the program staff will add trustworthiness to the data generated. And further, arguably, if decision-makers trust the data then both they themselves and their superiors will use that data. Regrettably, though, the involvement of both the program manager/and the evaluation staffs at the goal clarification stage of the might possibly create a struggle for power or withdrawal of support team-workmanship. The evaluation requirement that goals of the program activities be operationalized or transformed into behavioral or measureable terms tends to place the evaluation staff in a very threatening and, indeed, potentially destructive position. This is so because, among other things, the technique of operationalizing goals may put the decision-maker on the spot as being somewhat inadequate. In addition to the evaluation threat to the decision-maker personally, the operationalization of goals may produce a change of goal and, by convention, a change of goal intimates a change of program. If the change is minor it could prove to be unharmful. But if the change of goal differs markedly from the original intent, then it could be very destructive to the present scheme of program operations.

It seems to be well established that the evaluation activities should run concurrently with the program activities. In fact, prominent evaluators[38] appear to espouse the viewpoint that, ideally, the formative evaluator come to the program activity at the beginning of its initial phase and not only work rather closely with the program manager but also remain there at the program site throughout its entire operational life. We believe that the preceding paragraph provides the wherewithal for a case being made against the trend that calls for the evaluator's early arrival to the program site. At the very least, we feel that if the evaluator arrives at the program site either at its beginning or during the program's embryonic stages, then a case depicting the evaluator as a menace to can be made. At least one author seems to be in agreement with our interest in seeing a reversal of the present trend. Goodwin points out that in the case where the evaluator arrives at the program site with the decision-maker the "Evaluation is unfair and stifling to the fledging program just beginning."[39] Goodwin's position appears to be well grounded in fact.

At this point in the history of program evaluation, we are unable to produce adequate test data to support the theory that the benefit or practical utility of having the evaluator begin his "on-site" evaluation work simultaneously with the decision-maker at the outset of the program activity outweighs the harm and potential harm that the evaluator's mere presence at that particular point in time possibly creates for the decision-maker.

We believe that there is a built-in conflict between the formative stage of the program and formative evaluation as we have intimated above. In particular, we believe that the matter of having the evaluator and the decision-maker work concurrently at-the-job-site pits the evaluator in the position of being a menace by the nature

of the interdependence of his work on the decision-maker personally as a much needed source of information and authoritative knowledge. Whereas the decision-maker's interest is primarily in the day to day operations of the program activity and the responsibility for achieveing the program's goals ultimately, the evaluator's interest is primarily in establishing a pertinent list of behavioral components for the overall program so as to be able to provide interim feedback incident to the achievement of the program's ultimate goals. We submit, however, that the co-existence of a working relationship from the program's beginning through to its ending is well-nigh impossible. The bifurcated interwoven role of both decision-maker and evaluator appears to be unrealistic because of the dichotomous interest of the parties involved and the "watch-dog" phobia that (rightly or wrongly) attaches to the evaluative function. It is for the foregoing reasons that that we view the evaluator more as menace than a service to the need of the consumer via the decision-maker.

NOTES

1. Thomas E. Hutchinson, "The Operationalization of Fuzzy Concepts: A Methodology," (Amherst, Mass.: Center for Educational for Educational Research and Development, School of Education, University of Massachusetts, 1968), (Mimeographed).

2. In the early 1970's Phi Delta Kappa, Inc's Research Advisory Committee established a National Study Committee on Evaluation from among the prominent American evaluators of that period. Dr. Daniel I. Stufflebeam was the Chairman of that Committee.

3. Malcolm Provus, author of Discrepancy Evaluation, was a member of the Committee cited in n.2, above.

4. According to its publishers: Sage Publications' Evaluation: A Systematic Approach, by Peter H. Rossi, Howard E. Freeman, & Sonia R. Wright is the first textbook on program evaluation. It is a 1979 publication.

5. C. Robert Pace, "Evaluation Perspectives: 1968," (Los Angeles: Center for the Study of Evaluation, Graduate School of Education, University of California at Los Angeles, CSE Report No. 8, December, 1968).

6. The forementioned wide assortment of labels that are currently in use to describe evaluation activities.

7. Malcolm Provus, "Evaluation of Ongoing Programs in the Public School System," Educational Evaluation: New Roles, New Means, Sixty-eighth Yearbook of the National Society for the Study of Education, Part II (Chicago: University of Chicago Press, 1969), pp. 242-83.

8. Daniel I. Stufflebeam and others, Educational Evaluation and Decision Making, (Itasca, Ill.: F.E. Peacock Publishers,1971), p. 320.

9. Stufflebeam and others, pp. 218-235.

10. Robert E. Stake, "The Countenance of Educational Evaluation," Teachers College Record, 1967, 58, 523-540.

11. Malcolm L. Provus, Discrepancy Evaluation: For Educational Program Improvement and Assessment, (Berkeley, Calif.: McCutchan Publishing Corporation, 1974); Also Discrepancy Evaluation: A Practitioner's Guide, (Charlottesville, Va.: Evaluation Research Center, University of Virginia, 1976) by Diane K. Yavorsky.

12. The evaluation demand fostered by the Title I ESEA funding of 1965.

13. Robert Colvert, "A College Survey," unpublished study, Evaluation Research Center, University of Virginia, 1976.

14. California based regional laboratories (SWRL) and CSE of UCLA.

15. United States, Educational Resource Information Center--a clearinghouse of sorts.

16. Peter H. Rossi, Howard E. Freeman, and Sonia R. Wright, Evaluation: A Systematic Approach (Beverly HIlls, Calif.: Sage Publications, 1979).

17. Michael Scriven reacted to Lee J. Cronbach's ideas as expressed in "Course Improvement Through Evaluation," (Teachers College Record, 1963, 64, 672-683), and expanded upon them from the point of view of a philosopher.

18. William L. Goodwin, "Evaluation in Early Childhood Education," Preschool Education: A Handbook for the Training of Early Childhood Educators, eds. Ralph W. Colvin and Esther M. Zaffiro (New York: Springer, 1974), pp.201-52.

19. Goodwin, p. 221.

20. Goodwin, p. 204.

21. The reader's attention is called to the fact that allegedly 1979 marks the date of the first evaluation textbook, despite the fact that the well established need dates back to the mid-1960's.

22. The author's oral conversation with Thomas E. Hutchinson in 1970.

23. This piecemeal approach to the training of evaluators does not blemish our earlier claim that the vast majority of evaluation practitioners have had no formal training whatsoever.

24. See the first few statements in the paragraph following n.7, above.

25. Leon Jones, "Operationalization of Fuzzy Concepts: Methodological Research and Applications" (paper presented at the meeting of the American Educational Research Association, Chicago, April, 1972).

26. Jerome R. Sanders, What Are We Doing in Formal Evaluation Training Programs? U.S., Educational Resources Information Center, ERIC Document ED 109 238.

27. Leon Jones, "A Basic Evaluation Model" (uncopyrighted).

28. The OFC methodology by Thomas E. Hutchinson is used.

29. Leon Jones, "The Process of Decision-making and Its Relationship to Contemporary Educational Evaluation models" (Washington, D.C.: School of Education, Howard University, 1978), (Memeographed.)

30. Goodwin, p. 222.

31. See page 6 (above) beginning with the first full paragraph.

32. Same as note 1 (above).

33. For a practical application of the DEM, the interested reader is referred to Discrepancy Evaluation: A Practical Guide, by Dian K. Yavorsky as cited in note 11 (above).

34. The essence of the program design is the effective use of its two salient aspects; namely, its IPO's (Inputs (I), Process (P), and Outputs (O)) and its various levels (I, II, & III) of the design network. Here, as in note 33, the interested reader is referred to Discrepancy Evaluation: A Practical Guide, as cited in note 11 (above).

35. Yavorsky, p. 5.

36. See note 31 (above).

37. See note 1, (above).

38. See notes 2 & 3 (above).

39. Goodwin, "Objection #2," pp.204-05.